(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 223 550 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
17.07.2002 Bulletin 2002/29

(51) Int Cl.[7]: **G06T 5/00**

(21) Application number: **01203798.2**

(22) Date of filing: **08.10.2001**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **20.10.2000 US 692930**

(71) Applicant: **EASTMAN KODAK COMPANY**
Rochester, New York 14650 (US)

(72) Inventors:
- **Chen, Shoupu, Eastman Kodak Company**
  Rochester, New York 14650-2201 (US)
- **Ray, Lawrence A., Eastman Kodak Company**
  Rochester, New York 14650-2201 (US)

(74) Representative: **Parent, Yves et al**
KODAK INDUSTRIE,
Département Brevets,
CRT - Zone Industrielle
71102 Chalon-sur-Saône Cedex (FR)

(54) Method for detecting skin color in a digital image

(57) A method for detecting skin color in a digital image having pixels in an RGB color space generally includes the steps of performing statistical analysis of the digital color image to determine the mean RGB color values; then, if the mean value of any one of the colors is below a predetermined threshold, applying a transformation to the digital image to move skin colors in the image toward a predetermined region of the color space; and employing the transformed space to locate the skin color pixels in the digital color image. More specifically, if the mean value of any one of the colors is below a predetermined threshold, a non-linear transformation is applied to the digital image to move skin colors in the image toward a predetermined region of the color space. Then, depending on the preceding step, either the digital image or the transformed digital image is converted from the RGB space to a generalized RGB space to produce a gRGB digital image; skin color pixels are detected within the gRGB digital image; a first skin color image mask is formed based on the detected skin color pixels; a masked gRGB image is generated using the first skin color image mask; and finally the skin color image mask is employed to locate the skin color pixels in the digital color image.

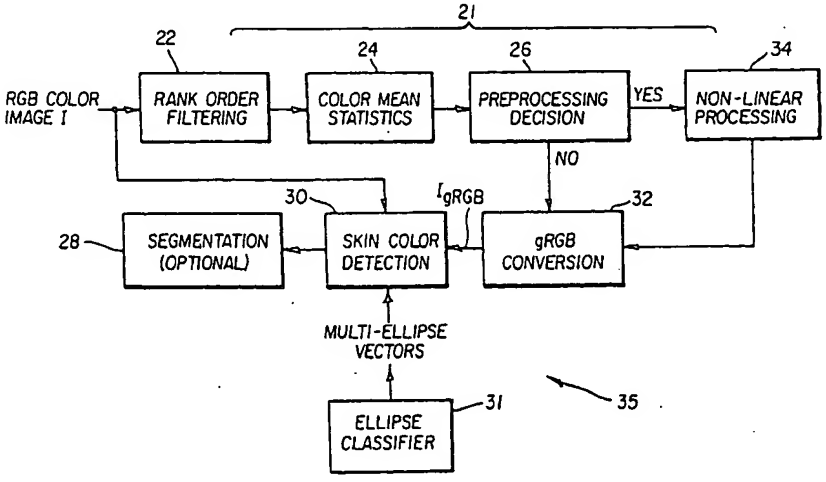

FIG. 2

EP 1 223 550 A2

Description

FIELD OF THE INVENTION

[0001] The present invention relates to digital image processing methods, and more particularly to such methods for detecting skin color in a digital image.

BACKGROUND OF THE INVENTION

[0002] In digital image processing it is often useful to detect the areas in the image that are skin color. This information is used, for example, to adjust the skin colors in the image to be pleasing. The location of skin color is also used in face detection and recognition algorithms, automatic image retrieval algorithms, and red-eye correction algorithms. For instance, U.S. Patent No. 4,203,671, issued May 20, 1980 to Takahashi et al., discloses a method of detecting skin color in an image by identifying pixels falling into an ellipsoid in red, green, blue color space or within an ellipse in two dimensional color space. The problem with this method is that it works well only when an image is properly balanced. For an over- or under-exposed image, the technique is not reliable. Furthermore, the technique does not work well for those skin colors that deviate from the chosen norm. For example, when the detection method is adjusted to detect light Caucasian skin, it fails to properly detect dark skin.

[0003] There is a need therefore for an improved skin color detection method that avoids the problems noted above.

SUMMARY OF THE INVENTION

[0004] The present invention is directed to overcoming one or more of the problems set forth above. Briefly summarized, according to one aspect of the present invention, a method for detecting skin color in a digital image having pixels in an RGB color space generally includes the steps of performing statistical analysis of the digital color image to determine the mean RGB color values; then, if the mean value of any one of the colors is below a predetermined threshold, applying a transformation to the digital image to move skin colors in the image toward a predetermined region of the color space; and employing the transformed space to locate the skin color pixels in the digital color image.

[0005] More specifically, if the mean value of any one of the colors is below a predetermined threshold, a non-linear transformation is applied to the digital image to move skin colors in the image toward a predetermined region of the color space. Then, depending on the preceding step, either the digital image or the transformed digital image is converted from the RGB space to a generalized RGB space to produce a gRGB digital image; skin color pixels are detected within the gRGB digital image; a first skin color image mask is formed based on the detected skin color pixels; a masked gRGB image is generated using the first skin color image mask; the masked gRGB image is converted to a hue image; possible blond hair color pixels are removed from the hue image to produce a modified hue image; a second skin color image mask is formed based on the skin color pixels in the modified hue image; if the second skin color image mask is smaller than the first skin color image mask by a predetermined amount, then the first skin color image mask is selected, otherwise, the second skin color image mask is selected; and finally the selected skin color image mask is employed to locate the skin color pixels in the digital color image.

[0006] The advantage of the invention is that it works well even when an image is not properly balanced, and furthermore works well for a variety of skin colors.

[0007] These and other aspects, objects, features and advantages of the present invention will be more clearly understood and appreciated from a review of the following detailed description of the preferred embodiments and appended claims, and by reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008]

Fig. 1 is a schematic diagram of an image processing system useful in practicing the present invention.
Fig. 2 is a flow chart illustrating the method stages of the present invention.
Fig. 3 is a detailed flow chart illustrating the skin color detection step shown in Fig. 2.
Fig. 4 shows an example of an ellipse classifier in two dimensional color space.
Fig. 5 shows a two dimensional color space having a plurality of skin color classifier regions, as used in the present invention.
Fig. 6 shows a two dimensional color space having skin color and blond hair color regions.
Fig. 7 is a schematic diagram describing the process used to train the ellipse classifier stage shown in Fig. 2.

DETAILED DESCRIPTION OF THE INVENTION

[0009] Because image processing algorithms and systems are well known, the present description will be directed in particular to algorithms and systems forming part of, or cooperating more directly with, the method in accordance with the present invention. Other aspects of such algorithms and systems, and hardware and/or software for producing and otherwise processing the image signals involved therewith, not specifically shown or described herein may be selected from such systems, algorithms, components and elements known in the art. In the following description, a preferred embodiment of the present invention will typically be implemented as software or a computer program. Those skilled in the art will readily recognize that the equivalent of such software may also be constructed in hardware. Given the method as described according to the invention in the following materials, software not specifically shown, suggested or described herein that is useful for implementation of the invention is conventional and within the ordinary skill in such arts.

[0010] Still further, as used herein, the computer program may be stored in a computer readable storage medium, which may comprise, for example; magnetic storage media such as a magnetic disk (such as a floppy disk or a hard drive) or magnetic tape; optical storage media such as an optical disc, optical tape, or machine readable bar code; solid state electronic storage devices such as random access memory (RAM), or read only memory (ROM); or any other physical device or medium employed to store a computer program.

[0011] Referring first to Fig. 1, an image processing system useful in practicing the present invention includes a color digital image source 10, such as a film scanner, digital camera, or digital image storage device such as a compact disc drive with a Picture CD. The digital image from the digital image source 10 is provided to an image processor 12, such as a programmed personal computer, or digital image processing workstation such as a Sun Sparc 20 workstation. The image processor 12 may be connected to a CRT display 14, and an operator interface such as a keyboard 16 and a mouse 18. The processed digital images are sent to an output device 20, such a hard copy printer, a long-term image storage device, a connection to another processor, or an image telecommunication device connected for example to the internet.

[0012] Referring to Fig. 2, the processing performed according to the present invention by the digital image processor 12 will be described. The digital image, expressed in a device independent RGB color space such as sRGB color space is first rank order filtered in a rank order filter step 22. Denote the input RGB image by $\mathbf{I} = \{\mathbf{C}_i\}$, where $i = 1,2,3$ for three color-channels, and $\mathbf{C}_i$ is a matrix with elements of $C_i(m, n)$, where $m = 0,...M - 1$ and $n = 0,...N - 1$. The rank-order filtering is defined as

$$C_i(m,n)\Big|_{C_i(m,n)\in\check{\mathbb{C}}_i} = median(C_i(s,t)\in\hat{\mathbb{C}}_i)\Big|_{s,m\in(0,M-1);n,t\in(0,N-1);s\neq m;t\neq n} \qquad (1)$$

where $\check{\mathbb{C}}_i$ is the set of $C_i(m,n)$ whose value exceeds a predefined threshold T_{Rank} while $\hat{\mathbb{C}}_i$ is the set of the remaining $C_i(m,n)$ of $\mathbf{I}$. This operation is similar to the traditional process of trimmed median filtering. Notice that the purpose of this rank-order filtering is not to improve the visual quality of the input image as traditional image processing does; rather, it is to reduce the influence of the pixels that have very high intensity values on the subsequent statistics gathering stage 24. For instance, in situations such as over exposure, or back lit exposure, pixels with values close to 255 will be altered or excluded in the mean value calculation. The resultant image of the rank-order filtering can be denoted by $\mathbf{I}_R = \{\hat{\mathbb{C}}_i, \check{\mathbb{C}}_i\}$, where the elements of $\hat{\mathbb{C}}_i$ have the original values, and the elements of $\check{\mathbb{C}}_i$ have the values computed using Equation 1. For simplicity, rewrite the rank-order filtered image as $\mathbf{I}_R = \{\bar{\mathbb{C}}_i\}$ where $\bar{\mathbb{C}}_i = \hat{\mathbb{C}}_i \cup \check{\mathbb{C}}_i$.

[0013] The next step is to compute the color mean-statistics in a statistics step 24 for each color channel using $\mathbf{I}_R$. This computation produces a set of mean values, that is, $\{m_i\} = mean(\mathbf{I}_R)$; where $m_i = mean(\bar{\mathbb{C}}_i)$. A pre-processing decision step 26 is made upon evaluating the mean statistics $\{m_i\}$. If there exits $m_i < T_M$, then the pre-processing decision is 'Yes', where the threshold T_M is an experimentally determined value; for example, for an 8-bit image, a value of 100 has been found to be acceptable.

[0014] If the pre-processing decision is 'Yes', then the process will go to a non-linear processing step 34. The input to the non-linear processing step is the original RGB image $\mathbf{I}$. The non-linear processor employed in this invention is color histogram equalization, which is an image processing technique well known to those of skill in this art and thoroughly described in the literature (e.g., see "Digital Image Processing", by Gonzalez and Woods, Addison-Wesley Publishing Company, 1992). The input image $\mathbf{I}$ is first converted to YIQ space to separate the luminance component and chromaticity components using the standard formula:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.569 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} \quad (2)$$

The histogram equalization process is applied to Y only, resulting in an equalized luminance Y'. The reverse mapping from the YIQ to RGB is done by

$$\begin{bmatrix} C_1' \\ C_2' \\ C_3' \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.273 & -0.647 \\ 1.000 & -1.104 & 1.701 \end{bmatrix} \begin{bmatrix} Y' \\ I \\ Q \end{bmatrix} \quad (3)$$

[0015] An important issue in color histogram equalization is saturation clipping. Remember that the equalization result Y' could cause C_i' to either exceed 255 or becomes less than 0. To prevent this from happening, a clipping procedure is used on C_i. The pseudo code for clipping saturation is as follows:

$$\begin{aligned} &if \quad C_1' > 255; \quad k_1 = 255/C_1' \\ &else \quad if \quad C_1' < 0; \quad C_1' = 0; \\ &if \quad C_2' > 255; \quad k_2 = 255/C_2' \\ &else \quad if \quad C_2' < 0; \quad C_2' = 0; \\ &if \quad C_3' > 255; \quad k_3 = 255/C_3' \\ &else \quad if \quad C_3' < 0; \quad C_3' = 0; \\ &k = \min(k_1, k_2, k_3) \\ &C_1' = kC_1', C_2' = kC_2', C_3' = kC_3' \end{aligned} \quad (4)$$

where k, k_1, k_2 and k_3 are initialized to zero.

[0016] If the pre-processing decision is "No", the original image I is sent to the gRGB conversion step 32.

[0017] In summary, the initialization stage 21 sends an image to a skin-color detection stage 35 depending on the pre-processing decision; either the original image **I** or the histogram-equalized image **I'** is sent to the gRGB conversion step 32. To unify the notation, define $\mathbf{I}_{RGB} = \{\mathbf{C}_i\}$, where i = 1,2,3 for three color-channels, and $\mathbf{C}_i$ is a matrix with elements of $C_i(m,n)$, where $m = 0,...M-1$ and $n = 0,...N-1$. Both the original image and the histogram-equalized image are denoted by $\mathbf{I}_{RGB}$ from now on.

[0018] Still referring to Figure 2, the input RGB image $\mathbf{I}_{RGB}$ is converted in the gRGB Conversion step 32 to a generalized RGB image, $\mathbf{I}_{gRGB}$, which is defined as $\mathbf{I}_{gRGB} = \{\mathbf{c}_i\}$, and $\mathbf{c}_i$ is a matrix with elements of $c_i(m,n)$, where m = 0,... $M-1$ and n = 0,...$N-1$. The conversion uses the formula:

$$c_j(m,n) = \frac{C_j(m,n)}{\sum_i C_i(m,n)} \mid j \in i = [1,2,3]; \quad m = [0,...M-1]; \quad n = [0,...N-1] \quad (5)$$

where C_i is the individual color channel (R, G, or B) of the input image. This conversion operation is not valid when

$$\sum_i C_i = 0,$$

, and the output will be set to zero. The resultant three new elements are linearly dependent, that is

$$\sum_j c_j = 1,$$

so that only two elements are needed to effectively form a new space (gRG plane) that is collapsed from three dimensions to two dimensions. In most cases, c_1 and c_2, that is, the generalized R and G, are used in skin color analysis and skin color detection. The skin colors in the image I_{gRGB} are detected in a skin color detection step 30 that receives multiple ellipse vectors from an ellipse classifier 31. The ellipse classifier 31 is trained on a large population of images beforehand, which will be discussed below with reference to Fig. 7. The detected skin colors may then be optionally segmented in a segmentation step 28.

[0019] The skin-color detection step 30 is shown in more detail in Figure 3. The generalized RGB image I_{gRGB} is first projected on to the gRG plane in the in a projection step 36. This projection produces a gRG image I_{gRG}. The projection is affected by simply removing the component c_3 from the gRGB image. The gRG image I_{gRG} with two components c_1 and c_2 is then passed to a first (I) skin color classifier 38 which receives a first ellipse vector (ellipse vector I) as shown in Figure 3. Figure 4 illustrates an example of an ellipse 58 in gRG color space 56 that is used to designate a region in which skin color pixels are located. The dark region 60 in Figure 4 is the actual skin-color pixel area. Parameters a, b, (o_1,o_2), and θ are the constants of the ellipse and define the ellipse vector produced by the ellipse classifier 31. Parameters *a* and *b* are the ellipse axes length, (o_1,o_2) is the center coordinates of the ellipse in the (c_1,c_2) space 56, and θ is the angle between the c_1 axis and the longer axis of the ellipse.

[0020] For the first skin-color classifier 38, define an intermediate skin-color mask $\mathbf{P}_I = \{p_I(i,j)\}_{M \times N}$, that is, an $M \times N$ matrix with elements $p(i,j)$. For each pixel of the generalized RG image I_{gRG}, define an evaluation function $E(c_1(i,j), c_2(i,j))$, where $i = 0,\cdots M-1$; $j = 0,\cdots N-1$, $c_1(i,j) \in [0,1]$; $c_2(i,j) \in [0,1]$. The classification is simply performed as:

$$p_I(i,j) = \begin{cases} 1 & if \quad E(c_1(i,j), c_2(i,j)) < 1 \\ 0 & else \end{cases} \qquad (6)$$

where "1" indicates a skin-color pixel and "0" a non-skin color pixel, and

$$E(c_1(i,j), c_2(i,j)) = x_{ij}^2/a^2 + y_{ij}^2/b^2 \qquad (7)$$

and

$$\begin{aligned} x_{ij} &= (c_1(i,j) - o_1)\cos(\theta - \pi/4) - (c_2(i,j) - o_2)\sin(\theta - \pi/4) \\ y_{ij} &= (c_1(i,j) - o_1)\sin(\theta - \pi/4) - (c_2(i,j) - o_2)\cos(\theta - \pi/4) \end{aligned} \qquad (8)$$

[0021] The intermediate skin-color mask, $\mathbf{P}_I$, is passed to an evaluation step 40. The evaluation procedure simply computes the ratio, γ_I, of the number of detected skin pixels to the image size. The evaluation step 40 has two branches. If $\gamma_I \geq T_\gamma$, that is, the evaluation result is 'Good', then the process branches to a skin color region masking step 42. In this case, the output from the first skin color classification step 38, $\mathbf{P}_I$, is used for masking the I_{gRGB} in the skin color region masking step 42. The output of the skin color region masking step 42 is the masked generalized RGB image: $\hat{I}_{gRGB} = I_{gRGB} \cap \mathbf{P}_I$. If $\gamma_I < T_\gamma$, then the evaluation result is 'Not good', and the process branches to a second skin color detection classification step 44. While it should be clear that T_γ may take on a range of different values depending on the application, a preferred value for T_γ is 0.12.

[0022] The structure and operation of the second (II) skin color classification step 44 is the same as that of the first skin color classification step 38 except it employs a different ellipse vector which determines the position, size and orientation of the ellipse in the (c_1,c_2) space. As shown in Figure 5, there are several possible ellipses E1-E7 that can be generated by the ellipse classifier 31. The idea of having multiple ellipses is to reduce false positives while providing a mechanism for detecting skin colors that deviate from a main skin color region, for example E1 in the (c_1,c_2) space 62. As an illustrative case, the use of only two such ellipses (I and II) is described with reference to Fig. 3, but it will be understood that a larger number of skin color classifiers and ellipses can be employed in the present invention.

[0023] There may be cases in which switching from the main skin color classifier (the first classifier step 38) to another classifier results in even fewer skin color pixels being detected. Therefore, the results from the skin color classification steps 38 and 44 are compared in an evaluation and selection step 46, even though the result from the first classifier 38 was not satisfied in the previous evaluation step 40. In a more general case, the selection decision is made as following:

$$\begin{aligned} &\gamma = \max_{\forall i}(\gamma_i) \\ &if \quad \gamma = \gamma_i \\ &then \quad \mathbf{P} = \mathbf{P}_i \end{aligned} \tag{9}$$

where **P** is the first skin-color mask defined as $\mathbf{P} = \{p(i,j)\}_{M\times N}$, and $\mathbf{P}_i$ is the intermediate skin-color mask from Skin color Classification I, II and so on if more than two classifiers are used. This first skin-color mask, **P**, will be used in the skin color region masking step 42.

[0024] If the result from the main skin color detector (the first classifier 38) is satisfied, then $\mathbf{P}_I$ will be directly used as the first skin color mask in the subsequent skin color region masking step 42. If not, then the process branches to another classifier, the second classifier 44 for example, and both the new result and the result from the main classifier (the first classifier 38) will be evaluated and the better (or the best) one will be selected as the first skin color mask to be sent to the masking step 42.

[0025] Notice that there are feedback paths 39 and 41 for the skin color classifiers 38 and 44 respectively in Figure 3. These feedback paths provide the detection result itself to the classifier for evaluation. The evaluation process is to compute the ratio of the number of detected skin pixels to the image size. If the ratio is smaller than a predefined threshold, the classifier ellipse is then expanded, for instance, by 30% or 40% more, to include more colors.

[0026] There are two outputs generated from the skin color region masking step 42. One is the masked gRGB image and the other one is the skin color mask **P** itself. These two outputs will be used in the subsequent optional blond-hair-color pixel removal that is described below.

[0027] Referring to Fig. 6, we have discovered that blond hair colored pixels 72 occur in the same region where the majority of skin colored pixels 71 reside, but next to the skin color region 70 that belong to the people having blond hairs. Most currently existing skin detection algorithms do not take the blond hair colored pixels into account. This can cause a problem for image understanding applications that are sensitive to the size of the area of skin color, such as redeye detection algorithms which rely on detecting the actual size of the classified skin color region. A technique for removal of blond hair colored pixels is described in the aforementioned copending Serial No. [our Docket No. 81321DMW], entitled "Method of Blond-Hair-Pixel Removal in Image Skin-Color Detection", which is incorporated herein by reference, and briefly summarized below.

[0028] It has been shown that difficulties arise when dealing with images having faces associated with blond hairs. In these cases, the conventional skin-color detection process fails to produce satisfied or desired results that would give help in redeye detection procedure. It is desirable to have blond-hair-color pixels removed from the masked skin-color image obtained by the steps described in the previous sections.

[0029] However, it is not a trivial task to parameterize the sub-regions such as the blond hair color region 72 and the skin color region 70 in the (c_1,c_2) space 68 so that the hair color can be separated from the face skin color. If the space dimension drops down to one, the separation of blond hair color pixels from the skin color pixels becomes fairly easy. This further reduction of dimension size is realized by converting the masked generalized RGB image $\hat{\mathbf{I}}_{gRGB}$ to a hue image $\mathbf{H} = \{h(m,n)\}_{M\times N}$ in a converting to hue image step 54. A typical hue conversion is performed as:

$$
\begin{aligned}
&if \quad (c_{\min} = \min_{i \in [1,2,3]}(c_i(m,n))) \neq (c_{\max} = \max_{i \in [1,2,3]}(c_i(m,n))) \quad do \\
&\quad if \quad c_1(m,n) = c_{\max} \quad do \quad h(m,n) = (c_2(m,n) - c_3(m,n))/(c_{\max} - c_{\min}) \\
&\quad elseif \quad c_2(m,n) = c_{\max} \quad do \quad h(m,n) = 2 + (c_3(m,n) - c_1(m,n))/(c_{\max} - c_{\min}) \\
&\quad elseif \quad c_3(m,n) = c_{\max} \quad do \quad h(m,n) = 4 + (c_1(m,n) - c_2(m,n))/(c_{\max} - c_{\min}); \\
&\qquad h(m,n) = h(m,n) * 60^{\circ}; \\
&\quad if \quad h(m,n) < 0.0 \quad do \quad h(m,n) = h(m,n) + 360^{\circ};
\end{aligned}
\tag{10}
$$

In a blond-hair color detection step 52, a predefined partition parameter T_H =15 is used to determine if an element $h(m,n)$ is a skin pixel or a blond hair pixel. A second mask, $\tilde{\mathbf{P}} = \{\tilde{p}(i,j)\}_{M \times N}$, is formed. If $h(m,n)$ is a skin pixel, then the corresponding element $\tilde{p}(m,n) = 1$, else $\tilde{p}(m, n) = 0$. In some cases, the blond-hair-color pixel removal may take away true skin-color pixels and the resultant second skin-color mask shrinks to an unusable small region. Therefore, the first skin-color mask **P** will be called back and the second the skin-color mask $\tilde{\mathbf{P}}$ is discarded. This action is performed in an evaluation and selection step 50 following the blond-hair color detection step 52 as shown in Figure 3. If the second skin color image mask is smaller than the first skin color mask by a predetermined amount, the first skin color image mask is selected, otherwise, the second skin color image mask is selected. The masked RGB color image $\tilde{\mathbf{I}}_{RGB}$ is the result of an AND operation of the selected skin color image mask and the original RGB color image $\mathbf{I}_{RGB}$. This operation is performed in a skin color region masking step 48.

[0030] Referring to Figure 7, a large image pool 74 containing over a thousand images from different sources is prepared for sample skin color patch collecting. It is not necessary to collect skin patches from all prepared images. A set of randomly selected images should be representative enough for the whole image pool, in theory and in practice as well. A set of uniformly distributed indices is used to retrieve the images from the image pool and form a set of random selected images 76. Sample skin color patches 78 are collected manually from the retrieved images. The collected skin color patches are then converted from the RGB space to the gRGB space 80 using Equation (5) above. The gRG data of the gRGB image is then evaluated in term of its distribution density 82 in the gRG plane 84, that is, the (c_1, c_2) space as described above. In practice, the distribution density is approximated with the histogram of the gRG data. The distribution density is further projected onto the (c_1, c_2) space after eliminating some of its elements whose height is less than 5% of the distribution density peak. The projected cluster 88 forms approximately an ellipse region in the gRG plane 86. The final step of the ellipse classifier training is to find the parameters for the ellipse region. The best-fit ellipse 92 in the gRG plane 90 is computed on the basis of moments. An ellipse is defined by its center (o_1, o_2), its orientation θ and its minor axis a and major axis b (see Figure 4). The center of the ellipse region is computed by

$$
\begin{aligned}
o_1 &= m_{10} / m_{00} \\
o_2 &= m_{01} / m_{00}
\end{aligned}
\tag{11}
$$

where the moments are computed as:

$$
m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c_1^p c_2^q f(c_1, c_2) dc_1 dc_2 \tag{12}
$$

where $f(c_1,c_2) = 1$, in this application. The orientation θ can be computed by determined the least moment of inertia:

$$
\theta = 0.5 \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{13}
$$

where the central moments are calculated as:

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(c_1 - o_1)^p (c_2 - o_2)^q f(c_1, c_2)dc_1 dc_2 \qquad (14)$$

and finally the length of minor and major axis can be computed as:

$$\begin{aligned} a &= 1.0623\left(A^3 / B\right)^{0.125} \\ b &= 1.0623\left(B^3 / A\right)^{0.125} \\ A &= \sum_{(c_1,c_2)\in ellipse_region}\left[(c_1 - o_1)\sin\theta - (c_2 - o_2)\cos\theta\right]^2 \\ B &= \sum_{(c_1,c_2)\in ellipse_region}\left[(c_1 - o_1)\cos\theta - (c_2 - o_2)\sin\theta\right]^2 \end{aligned} \qquad (15)$$

The above computation provides an initial set of ellipse vector $[a,b,o_1,o_2,\theta]$ 94. In practice, manual adjustment is needed to best fit the final ellipse to the gRG data cluster 88.

[0031] The subject matter of the present invention relates to digital image understanding technology, which is understood to mean technology that digitally processes a digital image to recognize and thereby assign useful meaning to human understandable objects, attributes or conditions and then to utilize the results obtained in the further processing of the digital image.

Claims

1. A method for detecting skin color in a digital image having pixels in an RGB color space, comprising the steps of:

   a) performing statistical analysis of the digital color image to determine the mean RGB color values;
   b) if the mean value of any one of the colors is below a predetermined threshold, applying a transformation to the digital image to move skin colors in the image toward a predetermined region of the color space; and
   c) employing the transformed space to locate the skin color pixels in the digital color image.

2. The method as claimed in claim 1 wherein step (c) further comprises the steps of:

   converting the transformed digital image from the RGB space to a generalized RGB space, to produce a gRGB digital image;
   detecting skin color pixels within the gRGB digital image;
   forming a first skin color image mask based on the detected skin color pixels; and
   employing the first skin color image mask to locate the skin color pixels in the digital color image.

3. The method as claimed in claim 2 wherein step (c) further comprises:

   forming a second skin color image mask based on the detected skin color pixels;
   selecting either the first skin color image mask or the second skin color image mask based on predetermined criteria; and
   employing the selected skin color image mask to locate the skin color pixels in the digital color image.

4. The method as claimed in claim 2 wherein step (c) further comprises the steps of:

   generating a masked gRGB image using the first skin color image mask;
   converting the masked gRGB image to a hue image;
   removing possible blond hair color pixels from the hue image to produce a modified hue image;
   forming a second skin color image mask based on the skin color pixels in the modified hue image;

if the second skin color image mask is smaller than the first skin color image mask by a predetermined amount, selecting the first skin color image mask, otherwise, selecting the second skin color image mask; and employing the selected skin color image mask to locate the skin color pixels in the digital color image.

5. A computer program product for performing the method as claimed in claim 1.

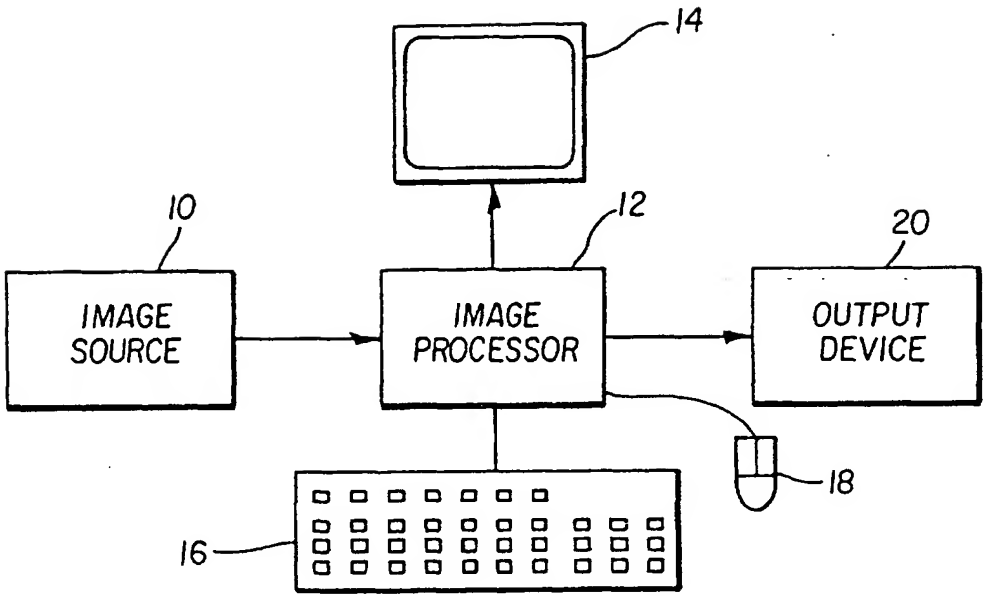

FIG. 1

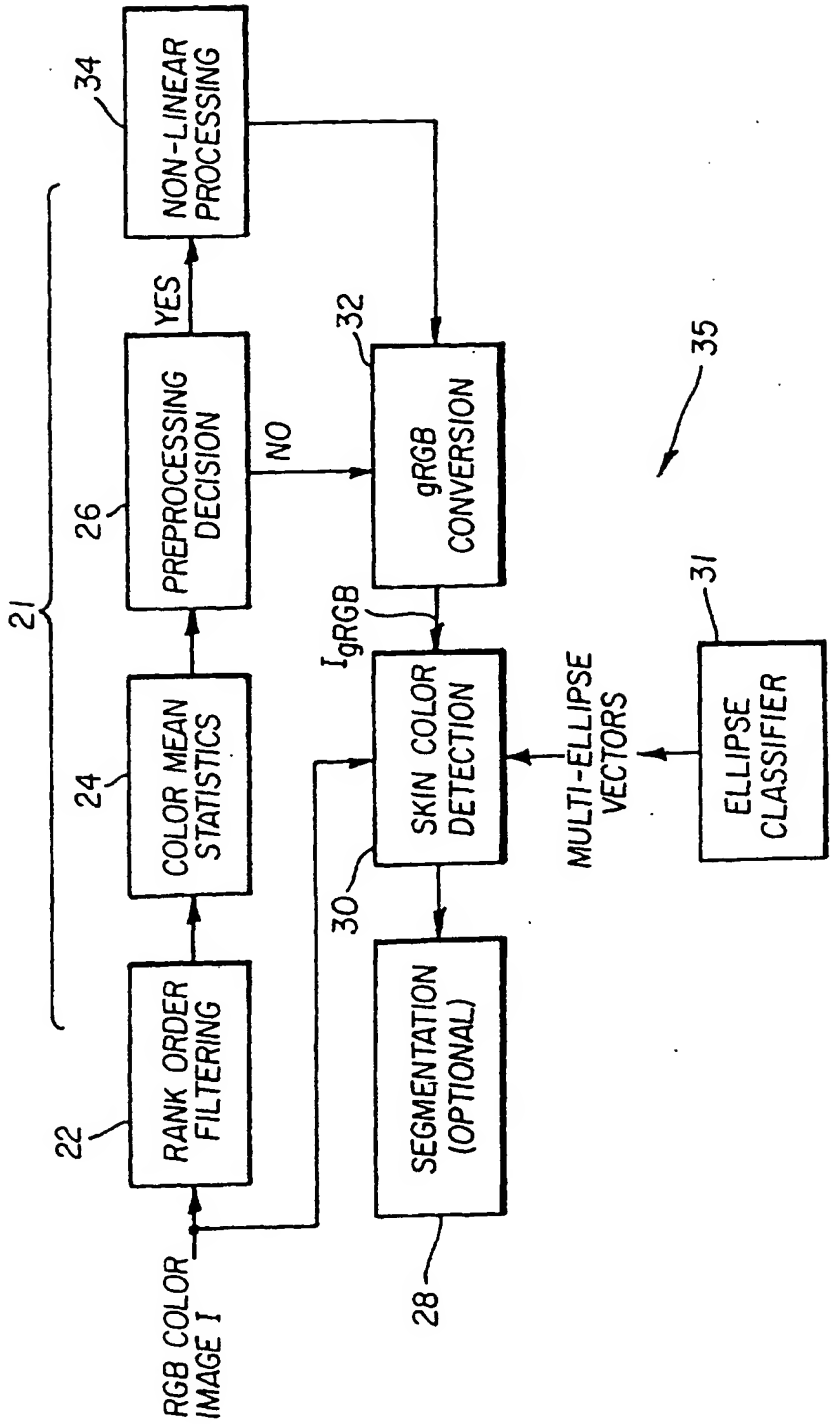
RGB COLOR IMAGE I
22
RANK ORDER FILTERING
24
COLOR MEAN STATISTICS
26
PREPROCESSING DECISION
YES
34
NON-LINEAR PROCESSING
NO
32
gRGB CONVERSION
I_{gRGB}
30
SKIN COLOR DETECTION
28
SEGMENTATION (OPTIONAL)
21
MULTI-ELLIPSE VECTORS
31
ELLIPSE CLASSIFIER
35

FIG. 2

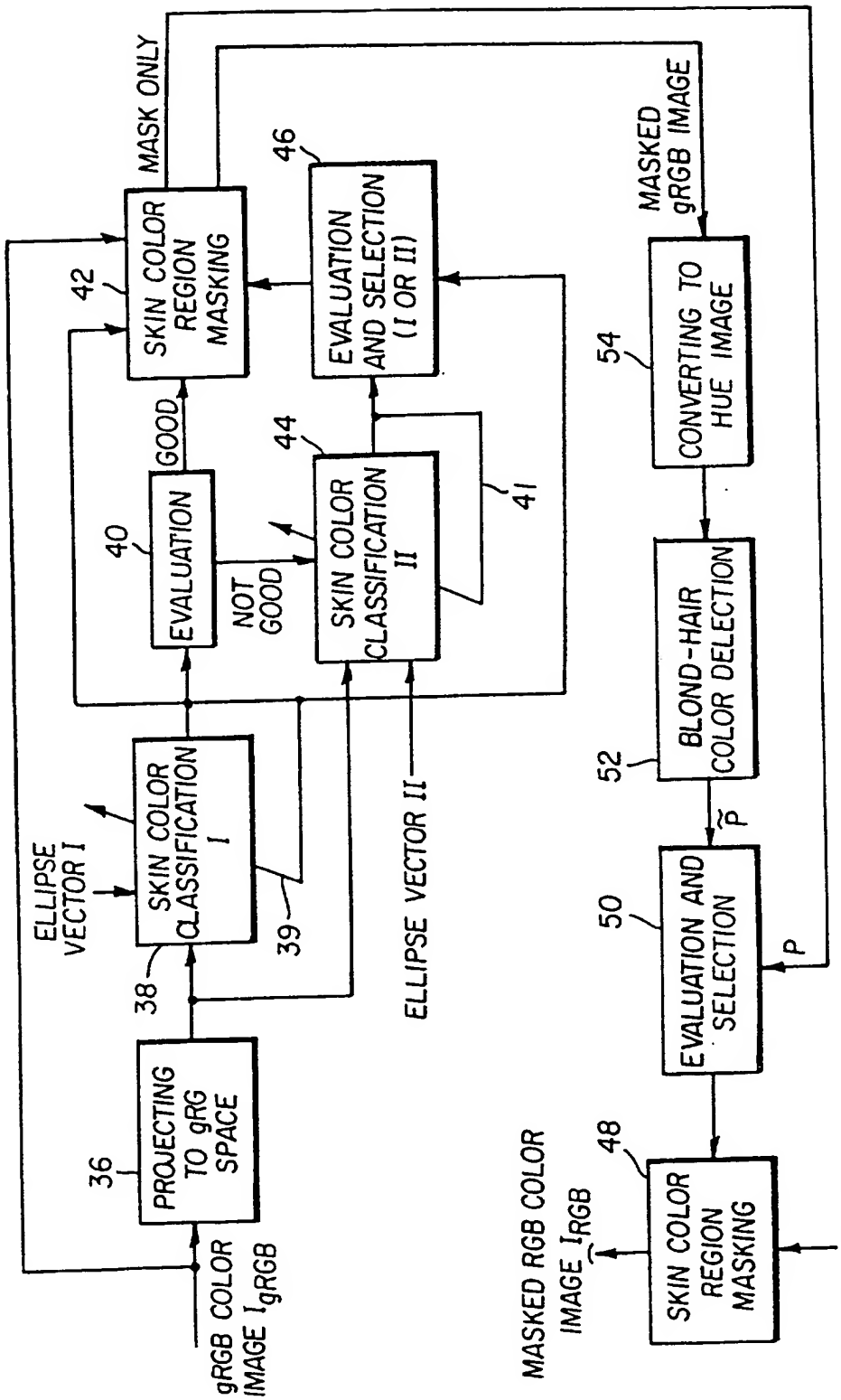
gRGB COLOR IMAGE I_{gRGB}
PROJECTING TO gRG SPACE
36
38
ELLIPSE VECTOR I
SKIN COLOR CLASSIFICATION I
39
ELLIPSE VECTOR II
EVALUATION
40
GOOD
NOT GOOD
SKIN COLOR CLASSIFICATION II
44
41
SKIN COLOR REGION MASKING
42
MASK ONLY
EVALUATION AND SELECTION (I OR II)
46
MASKED gRGB IMAGE
CONVERTING TO HUE IMAGE
54
BLOND-HAIR COLOR DELECTION
52
$\tilde{P}$
EVALUATION AND SELECTION
50
P
MASKED RGB COLOR IMAGE I_{RGB}
SKIN COLOR REGION MASKING
48
RGB COLOR IMAGE I

FIG. 3

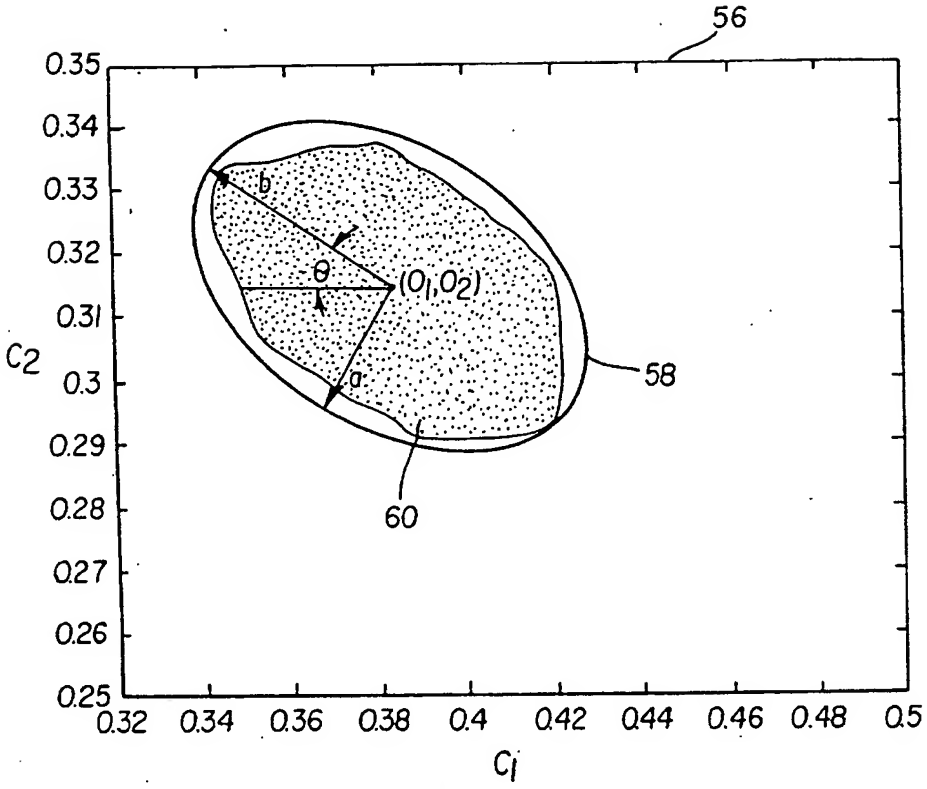
56
0.35
0.34
0.33
0.32
0.31
0.3
0.29
0.28
0.27
0.26
0.25
C2
0.32
0.34
0.36
0.38
0.4
0.42
0.44
0.46
0.48
0.5
C1
b
θ
(O1,O2)
a
58
60

FIG. 4

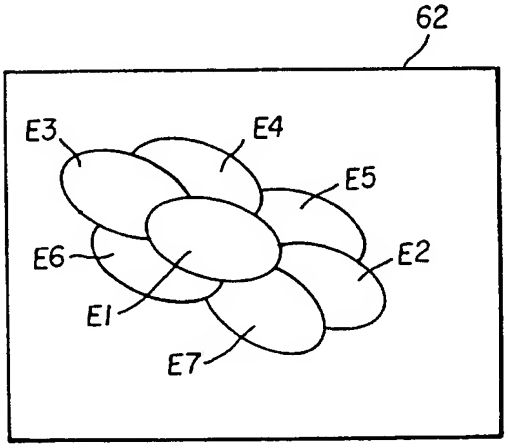
62
E3
E4
E5
E6
E2
E1
E7

FIG. 5

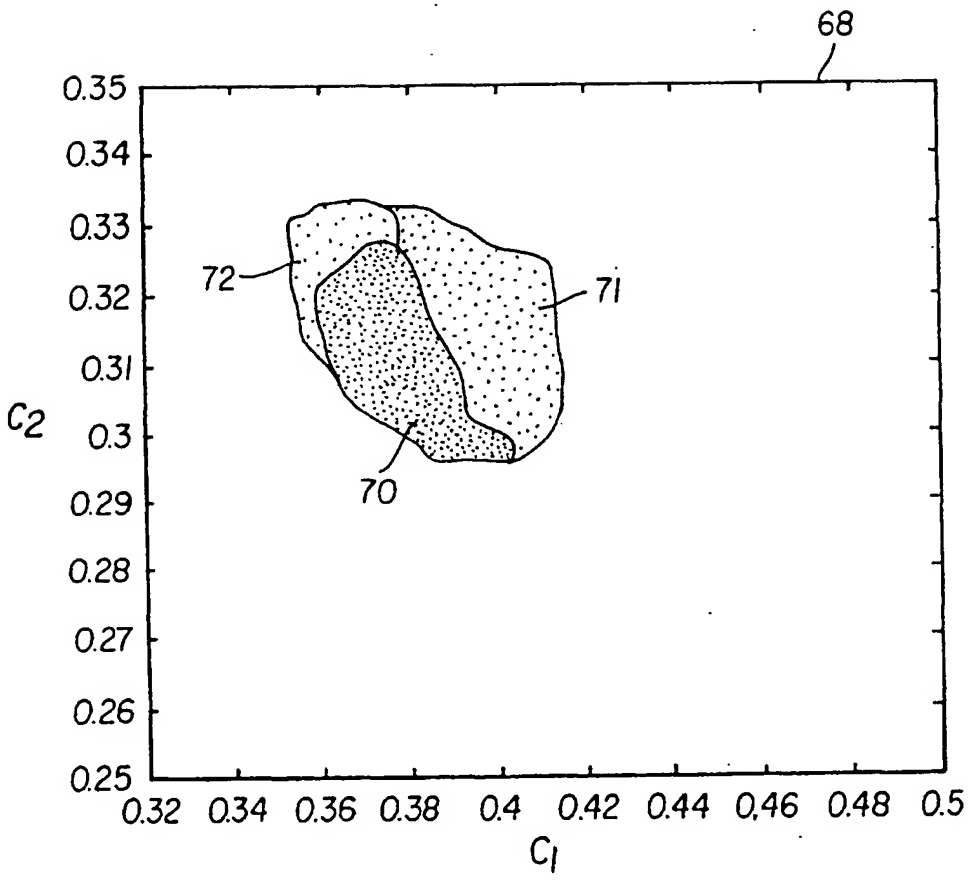
68
0.35
0.34
0.33
0.32
0.31
0.3
0.29
0.28
0.27
0.26
0.25
C2
72
71
70
0.32
0.34
0.36
0.38
0.4
0.42
0.44
0.46
0.48
0.5
C1

FIG. 6

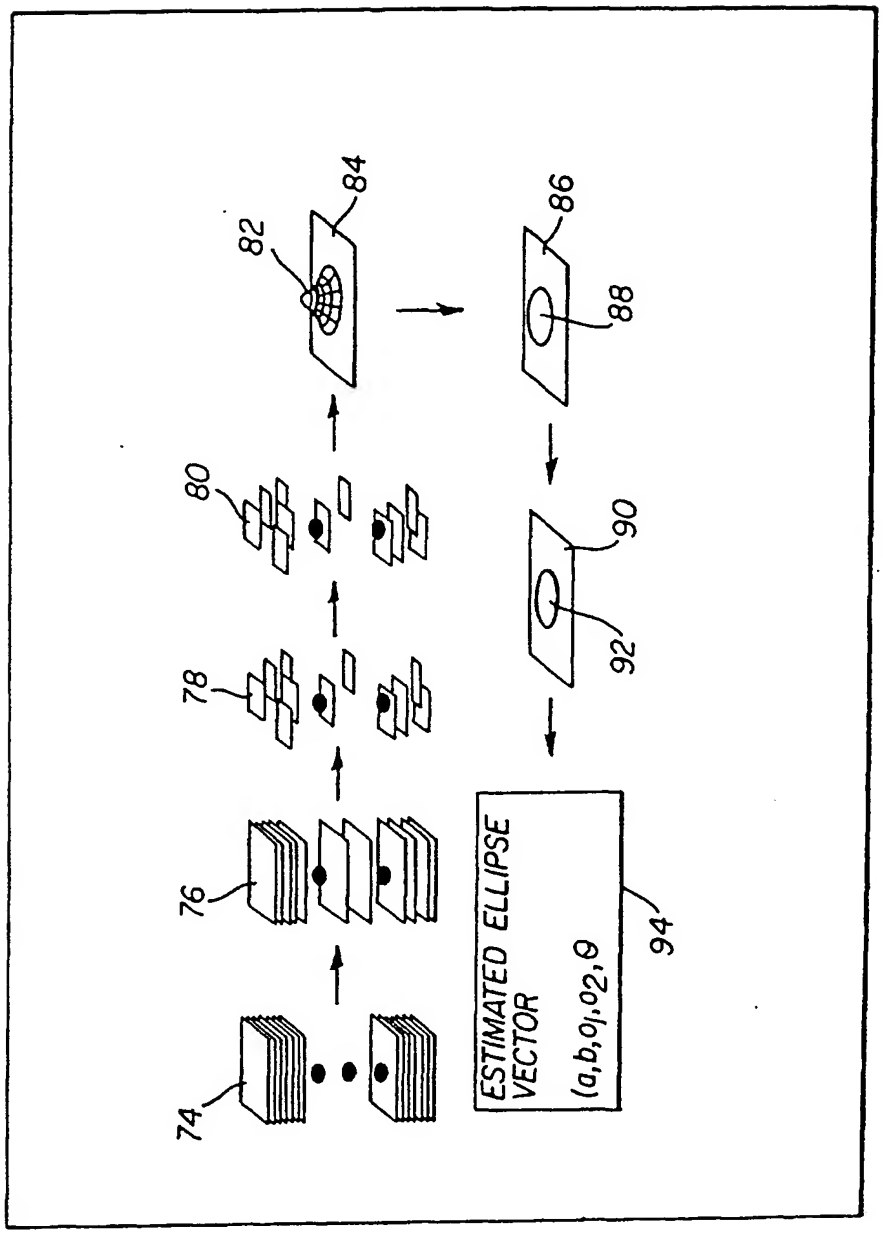
74
76
78
80
82
84
86
88
90
92
94
ESTIMATED ELLIPSE VECTOR
(a,b,o1,o2,θ

FIG. 7